## *Amendments*

***In the Specification:***

Please amend the specification as follows:

***In the Specification at page 1, line 2, immediately following the title please enter the following new sentence:***

This application claims priority under 35 U.S.C. § 119(e) to United States Provisional Application No. 60/144,351, filed July 20, 1999, and to United States Provisional Application No. 60/163,469, filed November 1, 1999.

***In the Specification at page 3, line 21 through page 4, line 5:***

Similarity analysis includes database search and alignment. Examples of public databases include those on the world wide web such as the DNA Database of Japan (DDBJ)(at ddbj.nig.ac.jp/); Genbank (at ncbi.nlm.nih.gov/web/Genbank/Index.html); and the European Molecular Biology Laboratory Nucleic Acid Sequence Database (EMBL) (at ebi.ac.uk/ebi_docs/embl_db.html). A number of different search algorithms have been developed, one example of which are the suite of programs referred to as BLAST programs. There are five implementations of BLAST, three designed for nucleotide sequences queries (BLASTN, BLASTX, and TBLASTX) and two designed for protein sequence queries (BLASTP and TBLASTN) (Coulson, *Trends in Biotechnology, 12*:76-80 (1994); Birren, *et al.*, *Genome Analysis, 1*:543-559 (1997)).

***In the Specification at page 8, lines 4-14:***

A characteristic feature of a large scale shotgun sequencing project is that the sequence data can be processed and assembled into contiguous sequences (contigs), which represent a reconstruction of the original genome sequence from the cloned fragments.

Likewise, individual Bacterial Artificial Chromosome (BAC) clones within a BAC library can be shot gun sequenced and these data can be assembled into contigs within each clone. Programs are available in the public domain that can analyze the sequence output and assemble the sequences into larger sequence regions representing contiguous sequences of the target genome. Examples of such programs can be found on the world wide web at, for example, genome.wustl.edu/gsc, sanger.ac.uk, and mbt.washington.edu. An example of sequence reading program is Phred (found on the world wide web at mbt.washington.edu). Phred reads DNA sequencer trace data, calls bases, assigns quality values to the bases, and writes the base calls and quality values to output files.

***In the Specification at page 8, line 15 through page 9, line 8:***

The process of assembling DNA sequence fragments generally involves three phases; the overlap phase, the layout phase and the multi-alignment, or consensus, phase. In the overlap phase, each fragment is compared against every other fragment to determine if they share a common subsequence, an indication that they were potentially sampled from overlapping stretches of the original DNA strand. Pairs of fragments are compared in two ways; 1) with both fragments in the same relative orientation, and 2) with one of the fragments having been reverse complemented. In the layout phase, a series of alternate assemblies or layouts of the fragments based on the pairwise overlaps is generated. A layout specifies the relative locations and orientations of the fragments with respect to each other and is typically visualized as an arrangement of overlapping directed lines, one for each fragment. The general criterion for the layout phase is to produce plausible assemblies of maximum likelihood. In this manner, it can be determined if there is more than one way to put the pieces together and if different solutions appear equally plausible.